Linear Algebra 2: Lecture 18

Irena Penev

Summer 2023

1 Algebraically closed fields

A field $\mathbb{F}$ is *algebraically closed* if every non-constant polynomial $p(x)$ with coefficients in $\mathbb{F}$ has a root in $\mathbb{F}$.

The Fundamental Theorem of Algebra. *Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.*

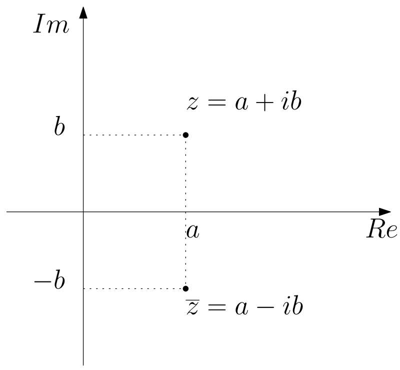
Note that the Fundamental Theorem of Algebra can be restated as saying that the field $\mathbb{C}$ is algebraically closed. Other fields that we have studied (namely, $\mathbb{R}$, $\mathbb{Q}$, and $\mathbb{Z}_p$ for a prime number p) are **not** algebraically closed. There exist algebraically closed fields other than $\mathbb{C}$, but we shall not study them in this course. Note, however, that if $\mathbb{F}$ is any algebraically closed field, and $p(x)$ is any polynomial of degree $n \geq 1$ with coefficients in $\mathbb{F}$, then there exist scalars $a, \alpha_1, \dots, \alpha_\ell \in \mathbb{F}$ such that $a \neq 0$ and such that $\alpha_1, \dots, \alpha_\ell$ are pairwise distinct, and positive integers $n_1, \dots, n_\ell$ satisfying $n_1 + \dots + n_\ell = n$, such that

$$p(x) = a(x - \alpha_1)^{n_1} \dots (x - \alpha_\ell)^{n_\ell}.$$

Here, a is the leading coefficient of $p(x)$, i.e. the coefficient in front of x^n. Scalars $\alpha_1, \dots, \alpha_\ell$ are the roots of $p(x)$ with *multiplicities* $n_1, \dots, n_\ell$, respectively. However, if $\mathbb{F}$ is a field that is **not** algebraically closed, then there exist non-constant polynomials $p(x)$ with coefficients in $\mathbb{F}$ that **cannot** be factored into linear factors as above.

In our study of eigenvalues and eigenvectors (to which we turn in section 2), we will need to factor polynomials at various stages. So, algebraically closed fields are of particular interest in the context of eigenvectors and eigenvalues.

We complete this section with a simple theorem (which might be familiar to you from high school). Recall that for a complex number $\mathbf{z} = a + ib$ (with $a, b \in \mathbb{R}$), we define the *complex conjugate* of z to be $\overline{z} = a - ib$. The complex conjugate of a complex number z is obtained by reflecting z across the real axis in the complex plane (see the picture below). Note that if z is a real number, then $\overline{z} = z$.



Theorem 1.1. *Let $p(x)$ be any polynomial with real coefficients, and let $z \in \mathbb{C}$. Then z is a root of $p(x)$ if and only if $\overline{z}$ is a root of $p(x)$.*

Proof. Set $p(x) = a_n x^n + \cdots + a_1 x + a_0$, where $a_0, a_1, \ldots, a_n \in \mathbb{R}$. Then we have the following sequence of equivalences:

$$\begin{aligned} p(z) = 0 \quad &\iff \quad \overline{p(z)} = \overline{0} \\ &\iff \quad \overline{a_n z^n + \cdots + a_1 z + a_0} = \overline{0} \\ &\overset{(*)}{\iff} \quad \overline{a_n}(\overline{z})^n + \cdots + \overline{a_1}(\overline{z}) = \overline{0} \\ &\overset{(**)}{\iff} \quad a_n(\overline{z})^n + \cdots + a_1\overline{z} + a_0 = 0 \\ &\iff \quad p(\overline{z}) = 0, \end{aligned}$$

where (*) follows the properties of the complex conjugate, and (**) follows from the fact that $a_0, a_1, \ldots, a_n$ and 0 are real numbers. □

Theorem 1.1 essentially states that if $p(x)$ is a polynomial with real coefficients, then its complex roots come in "conjugate pairs" (any real root that $p(x)$ may have is its own conjugate pair). It is not hard to show that the multiplicity of any complex root z of a polynomial $p(x)$ with real coefficients is equal to the multiplicity of the root $\overline{z}$, but we omit the details. However, this only works if $p(x)$ has **real** coefficients! If $p(x)$ has complex coefficients, then Theorem 1.1 does not apply.

2 Eigenvectors and eigenvalues

Given a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, we say that a scalar $\lambda \in \mathbb{F}$ is an *eigenvalue* of A if there exits a vector $\mathbf{v} \in \mathbb{F}^n \setminus \{\mathbf{0}\}$, called an *eigenvector* of A associated with the eigenvalue λ, such that

$$A\mathbf{v} = \lambda \mathbf{v}.$$

Thus, multiplying a matrix A by one of its eigenvectors has the effect of scaling that eigenvector, i.e. multiplying it by the associated eigenvalue. Note that, by definition, $\mathbf{0}$ is **not** an eigenvector of any matrix $A \in \mathbb{F}^{n\times n}$. This is because for any matrix $A \in \mathbb{F}^{n\times n}$ and any scalar λ, we have $A\mathbf{0} = \lambda\mathbf{0}$; if we allowed $\mathbf{0}$ to count as an eigenvector, then any scalar in $\mathbb{F}$ would be an eigenvalue, which would not be very interesting. However, the scalar 0 may possibly be an eigenvalue of a square matrix A.

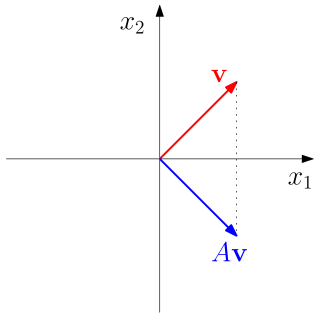
Later in this lecture, we will study how eigenvalues of a matrix can actually be computed (at least in some cases). First, let us take a look at some simple examples.

Example 2.1.

(a) Consider the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

in $\mathbb{R}^{2\times 2}$. *Note that* A *is the standard matrix of reflection about the* x_1*-axis in* $\mathbb{R}^2$.



The matrix A *has two eigenvalues:*

- $\lambda_1 = 1$ *is an eigenvalue of* A, *and* $\mathbf{v}_1 = \mathbf{e}_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$ *is an associated eigenvector, since*

$$A\mathbf{v}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \lambda_1\mathbf{v}_1;$$

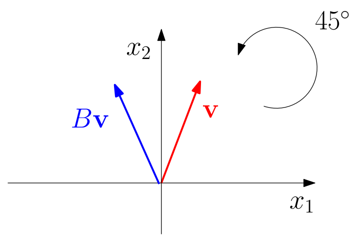
- $\lambda_2 = -1$ *is an eigenvalue of* A, *and* $\mathbf{v}_2 = \mathbf{e}_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$ *is an associated eigenvector, since*

$$A\mathbf{v}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \lambda_2\mathbf{v}_2.$$

(b) Consider the matrix

$$B = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

in $\mathbb{R}^{2\times 2}$. *Note that* B *is the standard matrix of counterclockwise rotation by* $45°$ *in* $\mathbb{R}^2$.



The matrix B *has no (real) eigenvalues, since it does not merely scale any non-zero vector in* $\mathbb{R}^2$.

(c) *Consider the matrix*

$$C = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

in $\mathbb{C}^{2\times 2}$. *(This is the same matrix as the one from part (b), except that now we consider it as a matrix in the vector space* $\mathbb{C}^{2\times 2}$*, which means that we are interested in its complex - not just real - eigenvalues.) Then:*

- $\lambda_1 = \frac{1+i}{\sqrt{2}}$ *is an eigenvalue of* C*, and* $\mathbf{v}_1 = \begin{bmatrix} i & 1 \end{bmatrix}^T$ *is an associated eigenvector, since*

$$C\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-1+i}{\sqrt{2}} \\ \frac{1+i}{\sqrt{2}} \end{bmatrix} = \lambda_1 \mathbf{v}_1;$$

- $\lambda_2 = \frac{1-i}{\sqrt{2}}$ *is an eigenvalue of* C*, and* $\mathbf{v}_2 = \begin{bmatrix} -i & 1 \end{bmatrix}^T$ *is an associated eigenvector, since*

$$C\mathbf{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -i \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-1-i}{\sqrt{2}} \\ \frac{1-i}{\sqrt{2}} \end{bmatrix} = \lambda_2 \mathbf{v}_2.$$

Note: The above shows that $\lambda_1 = \frac{1+i}{\sqrt{2}}$ *and* $\lambda_2 = \frac{1-i}{\sqrt{2}}$ *are indeed eigenvalues of* C*, but it is not at all obvious how those eigenvalues were computed. We will later see how eigenvalues and the corresponding eigenvalues can actually be computed. In this case of the matrix* C *above, this is worked out in detail in Example 2.8.*

(d) *For any field* $\mathbb{F}$ *and scalars* $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$*, consider the diagonal* *matrix*

$$D(\lambda_1, \lambda_2, \ldots, \lambda_n) := \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \mathbf{e}_1 & \lambda_2 \mathbf{e}_2 & \ldots & \lambda_n \mathbf{e}_n \end{bmatrix}.$$

Then $\lambda_1, \ldots, \lambda_n$ are all eigenvalues of $D(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Indeed, for each $i \in \{1, 2, \ldots, n\}$, the i-th standard vector $\mathbf{e}_i$ is an eigenvector of $D(\lambda_1, \lambda_2, \ldots, \lambda_n)$ associated with the eigenvalue λ_i, since

$$D(\lambda_1, \lambda_2, \ldots, \lambda_n)\mathbf{e}_i = \lambda_i \mathbf{e}_i.$$

Given a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ of A, the *eigenspace* of A associated with λ is the set

$$E_\lambda(A) := \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda \mathbf{v}\}.$$

So, the elements of $E_\lambda(A)$ are all the eigenvectors of A associated with λ, plus the zero vector.[1] When the matrix A is clear from context, we write just E_λ instead of $E_\lambda(A)$.

Theorem 2.2. *Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let λ be an eigenvalue of A. Then E_λ is a subspace of $\mathbb{F}^n$.*

Proof. We will apply Theorem 2.7 of Lecture Notes 6. Let us verify that the hypotheses of that theorem are satisfied.

1. We have that $A\mathbf{0} = \mathbf{0} = \lambda \mathbf{0}$, and so $\mathbf{0} \in E_\lambda$.
2. Fix $\mathbf{u}, \mathbf{v} \in E_\lambda$. Then

$$\begin{aligned} A(\mathbf{u} + \mathbf{v}) &= A\mathbf{u} + A\mathbf{v} \\ &= \lambda \mathbf{u} + \lambda \mathbf{v} \qquad \text{because } \mathbf{u}, \mathbf{v} \in E_\lambda \\ &= \lambda(\mathbf{u} + \mathbf{v}), \end{aligned}$$

and so $\mathbf{u} + \mathbf{v} \in E_\lambda$.

3. Fix $\mathbf{u} \in E_\lambda$ and $\alpha \in \mathbb{F}$. Then

$$\begin{aligned} A(\alpha \mathbf{u}) &= \alpha(A\mathbf{u}) \\ &= \alpha(\lambda \mathbf{u}) \qquad \text{because } \mathbf{u} \in E_\lambda \\ &= \lambda(\alpha \mathbf{u}), \end{aligned}$$

[1] We have that $\mathbf{0} \in E_\lambda(A)$ because $A\mathbf{0} = \lambda \mathbf{0}$.

and so $\alpha\mathbf{u} \in E_\lambda$.

We have now verified that E_λ satisfies the hypotheses of Theorem 2.7 of Lecture Notes 6, and so by that theorem, E_λ is indeed a subspace of $\mathbb{F}^n$. $\square$

For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times n}$, the *geometric multiplicity* of an eigenvalue λ of A is defined to be $\dim(E_\lambda)$.

2.1 Computing eigenvalues: the characteristic polynomial

Given a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times n}$, the *characteristic polynomial* of A is defined to be

$$p_A(\lambda) \quad := \quad \det(\lambda I_n - A).$$

The *characteristic equation* of A is the equation

$$\det(\lambda I_n - A) \quad = \quad 0.$$

So, the roots of the characteristic polynomial of A are precisely the solutions of the characteristic equation of A.

Example 2.3. *Compute the characteristic polynomial of the matrix*

$$A \quad = \quad \begin{bmatrix} 1 & -2 & 3 \\ -1 & 0 & 2 \\ 2 & -1 & -3 \end{bmatrix}$$

in $\mathbb{C}^{3\times 3}$.

Solution. The characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_3 - A) \\ &= \begin{vmatrix} \lambda - 1 & 2 & -3 \\ 1 & \lambda & -2 \\ -2 & 1 & \lambda + 3 \end{vmatrix} \\ &= \lambda^3 + 2\lambda^2 - 9\lambda - 3. \end{aligned}$$

$\square$

Remark: For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1. In some texts, the characteristic polynomial is defined to be $\det(A - \lambda I_n)$. But note that $\det(\lambda I_n - A) = (-1)^n \det(A - \lambda I_n)$, and so the polynomials $\det(\lambda I_n - A)$ and $\det(A - \lambda I_n)$ have exactly the same roots, with the same corresponding multiplicities, which is what we will actually care about when it comes to the characteristic polynomial. The main advantage of using $\det(\lambda I_n - A)$ rather than $\det(A - \lambda I_n)$ is that the former polynomial has leading coefficient 1, whereas the latter has leading coefficient $(-1)^n$, which is -1 if n is odd.

Theorem 2.4. *Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$, and let $\lambda_0 \in \mathbb{F}$. Then the following are equivalent:*

(i) λ_0 *is an eigenvalue of* A*;*

(ii) λ_0 *is a root of the characteristic polynomial of* A*, i.e.* $p_A(\lambda_0) = 0$*;*

(iii) λ_0 *is a solution of the characteristic equation of* A*, i.e.* $\det(\lambda_0 I_n - A) = 0$.

Proof. The fact that (ii) and (iii) are equivalent follows immediately from the definition of the characteristic polynomial and the characteristic equation of A. It remains to prove that (i) and (iii) are equivalent. For this, we have the following sequence of equivalences:

$$\begin{aligned}
\lambda_0 \text{ is an eigenvalue of } A \quad &\Longleftrightarrow \quad \exists \mathbf{v} \in \mathbb{F}^n \setminus \{\mathbf{0}\} \text{ s.t. } A\mathbf{v} = \lambda_0 \mathbf{v} \\
&\Longleftrightarrow \quad \exists \mathbf{v} \in \mathbb{F}^n \setminus \{\mathbf{0}\} \text{ s.t. } (\lambda_0 I_n - A)\mathbf{v} = \mathbf{0} \\
&\Longleftrightarrow \quad \begin{array}{l}\text{the equation } (\lambda_0 I_n - A)\mathbf{x} = \mathbf{0} \\ \text{has a non-trivial solution}\end{array} \\
&\overset{(*)}{\Longleftrightarrow} \quad \begin{array}{l}\text{the matrix } \lambda_0 I_n - A \\ \text{is non-invertible}\end{array} \\
&\overset{(**)}{\Longleftrightarrow} \quad \det(\lambda_0 I_n - A) = 0,
\end{aligned}$$

where (*) follows from Corollary 5.1 of Lecture Notes 4, and (**) follows from Theorem 5.1 from Lecture Notes 15. This proves that (i) and (iii) are equivalent, and we are done. □

For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n\times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is the largest integer k such that $(\lambda - \lambda_0)^k | p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.[2] Note that the sum of algebraic multiplicities of the matrix $A \in \mathbb{F}^{n\times n}$ is at most n; if the field $\mathbb{F}$ is algebraically closed, then the sum of algebraic multiplicities of A is exactly n.[3]

Theorem 2.5. *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.*

[2]In other words, k is the largest integer such that there exists some polynomial $q(\lambda)$ with coefficients in $\mathbb{F}$ such that $p_A(\lambda) = (\lambda - \lambda_0)^k q(\lambda)$.

[3]Indeed, if $\mathbb{F}$ is algebraically closed, then the characteristic polynomial $p_A(\lambda)$ can be written as a product of linear factors, and there are n of those factors. If $\mathbb{F}$ is not algebraically closed, we might or might not be able to factor $p_A(\lambda)$ in this way, which is why the sum of algebraic multiplicities of A is at most n (possibly strictly smaller than n).

Proof. Omitted. □

Schematically, for an eigenvalue λ of a matrix $A \in \mathbb{F}^{n\times n}$ (where $\mathbb{F}$ is an arbitrary field), Theorem 2.5 states that:

$$\text{geometric multiplicity of } \lambda \quad \leq \quad \text{algebraic multiplicity of } \lambda.$$

Example 2.6. *Consider the matrix*

$$A \quad = \quad \begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix}$$

in $\mathbb{C}^{3\times 3}$.

(a) Compute the characteristic polynomial $p_A(\lambda)$ *of the matrix* A.

(b) Find all the eigenvalues of A *and their algebraic multiplicities.*

(c) For each eigenvalue λ *of* A*, find a basis for the eigenspace* E_λ *and specify the geometric multiplicity of the eigenvalue* λ.

Solution. (a) The characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_3 - A) \\ &= \begin{vmatrix} \lambda-4 & 0 & 2 \\ -2 & \lambda-5 & -4 \\ 0 & 0 & \lambda-5 \end{vmatrix} \\ &\overset{(*)}{=} (\lambda-4)(\lambda-5)^2 \\ &= \lambda^3 - 14\lambda^2 + 65\lambda - 100, \end{aligned}$$

where the easiest way to obtain (*) is via Laplace expansion along the second column.

(**Remark:** We did not really need to expand in the last line. We only really care about the roots of the characteristic polynomial, and it is more convenient to have a form that is already factored.)

(b) From part (a), we see that A has two eigenvalues, namely, the eigenvalue $\lambda_1 = 4$ (with algebraic multiplicity 1), and the eigenvalue $\lambda_2 = 5$ (with algebraic multiplicity 2).

(c) For each $i \in \{1, 2\}$, the eigenspace E_{λ_i} is precisely the set of solutions of the equation $A\mathbf{x} = \lambda_i\mathbf{x}$, which is obviously equivalent to the equation

$$(\lambda_i I_3 - A)\mathbf{x} \quad = \quad \mathbf{0}.$$

For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_3 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -t/2 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} = \frac{t}{2} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So, $\left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \right\}$ is a basis of the eigespace E_{λ_1},[4] and we see that the eigenvalue $\lambda_1 = 4$ has geometric multiplicity 1.

For $\lambda_2 = 5$, we have that

$$\lambda_2 I_3 - A = \begin{bmatrix} \lambda_2 - 4 & 0 & 2 \\ -2 & \lambda_2 - 5 & -4 \\ 0 & 0 & \lambda_2 - 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 0 & -4 \\ 0 & 0 & 0 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_2 I_3 - A) = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_3 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -2t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{with } s, t \in \mathbb{C}.$$

So, $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of the eigenspace E_{λ_2}, and we see that the eigenvalue $\lambda_2 = 5$ has geometric multiplicity 2. □

---

[4] It is also true that $\left\{ \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} \right\}$ is a basis of E_{λ_1}. However, it is nicer to get integers (when possible).

Example 2.7. *Consider the matrix*

$$B = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

in $\mathbb{R}^{2\times 2}$.

> ***Remark:*** *This is the matrix from Example 2.1(b).*

(a) Compute the characteristic polynomial $p_B(\lambda)$ *of the matrix* B.

(b) Find all the (real) eigenvalues of B *and their algebraic multiplicities.*

> ***Remark:*** *Since we consider* B *to be a matrix in* $\mathbb{R}^{2\times 2}$, *we need to look for* ***real*** *eigenvalues only.*

(c) For each eigenvalue λ *of* B, *find a basis for the eigenspace* E_λ *and specify the geometric multiplicity of the eigenvalue* λ.

Solution. (a) The characteristic polynomial of B is:

$$\begin{aligned} p_B(\lambda) &= \det(\lambda I_2 - B) \\ &= \begin{vmatrix} \lambda - \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \lambda - \frac{1}{\sqrt{2}} \end{vmatrix} \\ &= (\lambda - \tfrac{1}{\sqrt{2}})(\lambda - \tfrac{1}{\sqrt{2}}) - \tfrac{1}{\sqrt{2}}(-\tfrac{1}{\sqrt{2}}) \\ &= \lambda^2 - \sqrt{2}\lambda + 1. \end{aligned}$$

(b,c) We need to find any real roots that the polynomial $p_B(\lambda)$ may have, i.e. any real solutions that the quadratic equation

$$\lambda^2 - \sqrt{2}\lambda + 1 = 0$$

may have. The discriminant of this quadratic equation is $(-\sqrt{2})^2 - 4 \cdot 1 \cdot 1 = -2 < 0$, and it follows that the equation has no real solutions. Therefore, B has no real eigenvalues. □

Example 2.8. *Consider the matrix*

$$C = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

in $\mathbb{C}^{2\times 2}$.

Remark: This is the matrix from Example 2.1(c). It is the same as the matrix B from Example 2.7, but this time, we consider the matrix to be in $\mathbb{C}^{2\times 2}$.

(a) Compute the characteristic polynomial $p_C(\lambda)$ of the matrix C.

(b) Find all the eigenvalues of C and their algebraic multiplicities.

> ***Remark:** Since we consider C to be a matrix in $\mathbb{C}^{2\times 2}$, we need to look for **complex** eigenvalues. (Note that all real numbers are complex! So, if our eigenvalues ended up being real, they would still count as complex eigenvalues.)*

(c) For each eigenvalue λ of C, find a basis for the eigenspace E_λ and specify the geometric multiplicity of the eigenvalue λ.

Solution. (a) The characteristic polynomial of C is the same as the characteristic polynomial of the matrix B from Example 2.7, since the two characteristic polynomials are computed in exactly the same way. Indeed,

$$\begin{aligned} p_C(\lambda) &= \det(\lambda I_2 - C) \\ &= \begin{vmatrix} \lambda - \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \lambda - \frac{1}{\sqrt{2}} \end{vmatrix} \\ &= (\lambda - \tfrac{1}{\sqrt{2}})(\lambda - \tfrac{1}{\sqrt{2}}) - \tfrac{1}{\sqrt{2}}(-\tfrac{1}{\sqrt{2}}) \\ &= \lambda^2 - \sqrt{2}\lambda + 1. \end{aligned}$$

(b) We need to find the (complex) roots of the characteristic polynomial $p_C(\lambda)$, together with their algebraic multiplicities. The quadratic equation

$$\underbrace{\lambda^2 - \sqrt{2}\lambda + 1}_{=p_C(\lambda)} = 0$$

has solutions

$$\lambda_{1,2} = \frac{-(-\sqrt{2}) \pm \sqrt{(-\sqrt{2})^2 - 4\cdot 1\cdot 1}}{2\cdot 1} = \frac{\sqrt{2} \pm \sqrt{-2}}{2} = \frac{1 \pm i}{\sqrt{2}},$$

that is,

$$\lambda_1 = \frac{1+i}{\sqrt{2}} \quad \text{and} \quad \lambda_2 = \frac{1-i}{\sqrt{2}}.$$

Complex numbers λ_1 and λ_2 are the eigenvalues of the matrix C, and they each have algebraic multiplicity 1, since $p_C(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)$.

(c) For each $i \in \{1, 2\}$, the eigenspace E_{λ_i} is precisely the set of solutions of the equation $C\mathbf{x} = \lambda_i \mathbf{x}$, which is obviously equivalent to the equation

$$(\lambda_i I_2 - C)\mathbf{x} = \mathbf{0}.$$

For $\lambda_1 = \frac{1+i}{\sqrt{2}}$, we have that

$$\lambda_1 I_2 - C = \begin{bmatrix} \lambda_1 - \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \lambda_1 - \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_1 I_2 - C) = \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_2 - C)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} it \\ t \end{bmatrix} = t \begin{bmatrix} i \\ 1 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So, $\left\{ \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}$ is a basis of the eigenspace E_{λ_1}, and we see that the eigenvalue $\lambda_1 = \frac{1+i}{\sqrt{2}}$ has geometric multiplicity 1.

For $\lambda_2 = \frac{1-i}{\sqrt{2}}$, we have that

$$\lambda_2 I_2 - C = \begin{bmatrix} \lambda_2 - \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \lambda_2 - \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_2 I_2 - C) = \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_2 I_2 - C)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -it \\ t \end{bmatrix} = t \begin{bmatrix} -i \\ 1 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So, $\left\{ \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\}$ is a basis of the eigenspace E_{λ_2}, and we see that the eigenvalue $\lambda_2 = \frac{1-i}{\sqrt{2}}$ has geometric multiplicity 1. □

Proposition 2.9. *Let $\mathbb{F}$ be a field, and let A be a triangular matrix in $\mathbb{F}^{n\times n}$. Then the eigenvalues of A are precisely the entries of A on its main diagonal, and moreover, the algebraic multiplicity of each eigenvalue is precisely the number of times that it appears on the main diagonal of A.*[5]

Proof. Set $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$. Since A is triangular, so is the matrix $\lambda I_n - A$; so, the determinant of $\lambda I_n - A$ can be computed simply by multiplying its entries on the main diagonal. It follows that the characteristic polynomial of A is

$$p_A(\lambda) = \det(\lambda I_n - A) = (\lambda - a_{1,1})(\lambda - a_{2,2})\dots(\lambda - a_{n,n}),$$

and the result follows. □

[5]However, the geometric multiplicity may possibly be smaller, as Example 2.10 shows.

Example 2.10. *Consider the matrix*

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

in $\mathbb{C}^{5\times 5}$.

(a) Compute the characteristic polynomial $p_A(\lambda)$ *of the matrix* A.

(b) Find all the eigenvalues of A *and their algebraic multiplicities.*

(c) For each eigenvalue λ *of* A*, find a basis for the eigenspace* E_λ *and specify the geometric multiplicity of the eigenvalue* λ.

Solution. (a) The matrix A is upper triangular, and so its characteristic polynomial is

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_5 - A) \\ &= \begin{vmatrix} \lambda-1 & -2 & 0 & 0 & 0 \\ 0 & \lambda-2 & 0 & 0 & 0 \\ 0 & 0 & \lambda-1 & -1 & -3 \\ 0 & 0 & 0 & \lambda-3 & -3 \\ 0 & 0 & 0 & 0 & \lambda-3 \end{vmatrix} \\ &= (\lambda-1)^2(\lambda-2)(\lambda-3)^2. \end{aligned}$$

(b) We see from part (a) that A has three eigenvalues, namely, $\lambda_1 = 1$ (with algebraic multiplicity 2), $\lambda_2 = 2$ (with algebraic multiplicity 1), and $\lambda = 3$ (with algebraic multiplicity 2).[6]

(c) For each $i \in \{1,2,3\}$, the eigenspace E_{λ_i} is precisely the set of solutions of the equation $A\mathbf{x} = \lambda_i\mathbf{x}$, which is obviously equivalent to the equation

$$(\lambda_i I_5 - A)\mathbf{x} = \mathbf{0}.$$

For $\lambda_1 = 1$, we have that

$$\lambda_1 I_5 - A = \begin{bmatrix} 0 & -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -3 \\ 0 & 0 & 0 & -2 & -3 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix},$$

[6] We could also have obtained the same answer by noticing that A is triangular, and that 1 appears twice on the main diagonal of A, 2 appears once, and 3 appears twice.

and that

$$\text{RREF}(\lambda_1 I_5 - A) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_5 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} s \\ 0 \\ t \\ 0 \\ 0 \end{bmatrix} = s \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \text{with } s, t \in \mathbb{C}.$$

So,

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

is a basis of the eigenspace E_{λ_1}, and we see that the eigenvalue $\lambda_1 = 1$ has geometric multiplicity 2.

For $\lambda_2 = 2$, we have that

$$\lambda_2 I_5 - A = \begin{bmatrix} 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & -1 & -3 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_2 I_5 - A) = \begin{bmatrix} 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Consequently, the general solution of the equation $(\lambda_2 I_5 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} 2t \\ t \\ 0 \\ 0 \\ 0 \end{bmatrix} = t \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So,

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

is a basis of the eigenspace E_{λ_2}, and we see that the eigenvalue $\lambda_2 = 2$ has geometric multiplicity 1.

For $\lambda_3 = 3$, we have that

$$\lambda_3 I_5 - A = \begin{bmatrix} 2 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -1 & -3 \\ 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and that

$$\mathrm{RREF}(\lambda_3 I_5 - A) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Consequently, the general solution of the equation $(\lambda_3 I_5 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ \frac{t}{2} \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} = \frac{t}{2} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So,

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} \right\}$$

is a basis of the eigenspace E_{λ_3}, and we see that the eigenvalue $\lambda_3 = 3$ has geometric multiplicity 1. □

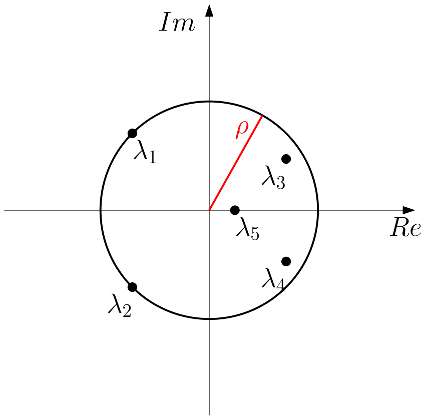
2.2 The spectrum of a matrix

Given a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, the *spectrum* of A is the multiset of all eigenvalues of A, and the number of times that each eigenvalue appears in the spectrum is precisely equal to the algebraic multiplicity of that eigenvalue.

For example, if $A \in \mathbb{C}^{5\times 5}$ has eigenvalues 1 (with algeraic multiplicity 1), $1+i$ (with algebraic multiplicity 2), and $1-i$ (with algebraic multiplicity 2), then the spectrum of A is $\{1, 1+i, 1+i, 1-i, 1-i\}$.

For a matrix $A \in \mathbb{C}^{n\times n}$, the *spectral radius* of A, denoted by $\rho(A)$, is the maximum absolute value of any eigenvalue of A. For example, if the spectrum of a matrix $A \in \mathbb{C}^{5\times 5}$ is $\{1, 1+i, 1+i, 1-i, 1-i\}$, then the spectral radius of A is $\rho(A) = \max\{|1|, |1+i|, |1+i|, |1-i|, |1-i|\} = \sqrt{2}$.[7]

In view of Theorems 1.1 and 2.4, we can visualize the complex eigenvalues of an $n \times n$ matrix A with **real** entries (however, we consider A to be a matrix in the vector space $\mathbb{C}^{n\times n}$, so that it can have complex eigenvalues). Its characteristic polynomial $p_A(\lambda)$ is of degree n and has real coefficients. By Theorem 1.1, the roots of this polynomial come in conjugate pairs,[8] and moreover, by Theorem 2.4, those roots are precisely the eigenvalues of A. The eigenvalues all lie in the complex plane, in the disk centered at the origin and with radius $\rho(A)$, and they are symmetric about the real axis. Visually, the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ of a matrix $A \in \mathbb{C}^{5\times 5}$ with real entries might look as in the picture below.



Recall that the *trace* of a square matrix $A = \left[\ a_{i,j}\ \right]_{n\times n}$ with entries in some field $\mathbb{F}$ is defined to be $\mathrm{trace}(A) := \sum_{i=1}^{n} a_{i,i}$, i.e. the trace of A is the sum of entries on the main diagonal of A.

Theorem 2.11. *Let $\mathbb{F}$ be a field, let $A = \left[\ a_{i,j}\ \right]$ be a matrix in $\mathbb{F}^{n\times n}$, and assume that $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A. Then*

(a) $det(A) = \lambda_1 \dots \lambda_n$;

(b) $trace(A) = \lambda_1 + \dots + \lambda_n$.

Remark: Theorem 2.11 only applies if the spectrum of the matrix $A \in \mathbb{F}^{n\times n}$ contains n eigenvalues (including algebraic multiplicities)! This will always happen if $\mathbb{F}$ is algebraically closed, but need not happen otherwise.

[7]Indeed, $|1| = 1$, $|1+i| = \sqrt{2}$, and $|1-i| = \sqrt{2}$. So, $\rho(A) = \sqrt{2}$.

[8]Each real root is its own conjugate pair.

Proof. By definition, we have that $p_A(\lambda) = \det(\lambda I_n - A)$. On the other hand, since $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A (and A is an $n \times n$ matrix), we see that $p_A(\lambda) = (\lambda - \lambda_1) \dots (\lambda - \lambda_n)$.

(a) By setting $\lambda = 0$, we obtain

$$p_A(0) \quad = \quad (0 - \lambda_1) \dots (0 - \lambda_n) \quad = \quad (-1)^n \lambda_1 \dots \lambda_n.$$

On the other hand, we have that

$$p_A(0) \quad = \quad \det(0I_n - A) \quad = \det(-A) \quad \overset{(*)}{=} \quad (-1)^n \det(A),$$

where (*) was obtained by iteratively multiplying the rows of $-A$ by -1 in order to obtain A in the end. We now deduce that $(-1)^n \lambda_1 \dots \lambda_n = (-1)^n \det(A)$, and it follows that $\det(A) = \lambda_1 \dots \lambda_n$.

(b) We will compute the coefficient in front of λ^{n-1} in the characteristic polynomial $p_A(\lambda)$ in two ways.

First, since $p_A(\lambda) = (\lambda - \lambda_1) \dots (\lambda - \lambda_n)$, it is clear that the coefficient in front of λ^{n-1} is $-\lambda_1 - \dots - \lambda_n$.

On the other hand, we have that

$$p_A(\lambda) \quad = \quad \det(\lambda I_n - A) \quad = \quad \begin{vmatrix} \lambda - a_{1,1} & -a_{1,2} & \dots & -a_{1,n} \\ -a_{2,1} & \lambda - a_{2,2} & \dots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \dots & \lambda - a_{n,n} \end{vmatrix}.$$

We now use the definition of the determinant: the only permutation $\sigma \in S_n$ that produces a polynomial with λ^{n-1} appearing with it (with a possibly non-zero coefficient) is the identity permutation $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$,[9] and clearly, this permutation is even, i.e. has sign 1. So, the coefficient in front of λ^{n-1} in $p_A(\lambda)$ is equal to the coefficient of λ^{n-1} in the product $(\lambda - a_{1,1})(\lambda - a_{2,2}) \dots (\lambda - a_{n,n})$, which is precisely $-a_{1,1} - a_{2,2} - \dots - a_{n,n} = -\text{trace}(A)$.

We have now computed the coefficient in front of λ^{n-1} in the polynomial $p_A(\lambda)$ in two ways: we got $-\lambda_1 - \dots - \lambda_n$ the first time, and we got $-\text{trace}(A)$ the second time. So, $-\lambda_1 - \dots - \lambda_n = -\text{trace}(A)$, and it follows that $\text{trace}(A) = \lambda_1 + \dots + \lambda_n$. □

2.3 More about eigenspaces

Theorem 2.12. *Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let $\lambda_1, \dots, \lambda_k$ be pairwise distinct eigenvalues of A. For each $i \in \{1, \dots, k\}$, let $\mathcal{B}_{\lambda_i} = \{\mathbf{v}_1^{\lambda_i}, \dots, \mathbf{v}_{t_i}^{\lambda_i}\}$ be a basis of the eigenspace E_{λ_i}. Then*

$$\{\mathbf{v}_1^{\lambda_1}, \dots, \mathbf{v}_{t_1}^{\lambda_1}, \mathbf{v}_1^{\lambda_2}, \dots, \mathbf{v}_{t_2}^{\lambda_2}, \dots, \mathbf{v}_1^{\lambda_k}, \dots, \mathbf{v}_{t_k}^{\lambda_k}\}$$

[9] Note that the identity permutation encodes the selection of the entire main diagonal.

is a linearly independent set of vectors in $\mathbb{F}^n$.

Proof. We begin by proving a claim.

> **Claim.** For all $\mathbf{v}_1 \in E_{\lambda_1} \setminus \{\mathbf{0}\}, \dots, \mathbf{v}_k \in E_{\lambda_k} \setminus \{\mathbf{0}\}$, the set $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent.

Proof of the Claim. Suppose otherwise, and consider an inclusion-wise minimal subset of $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ that is linearly dependent. After possibly permuting the order of our eigenvalues, we may assume that there exists some $\ell \in \{1, \dots, k\}$ such that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is an inclusion-wise minimal linearly dependent set, i.e. it is linearly dependent, but all its proper subsets are linearly independent. Since $\mathbf{v}_1 \neq \mathbf{0}$, we see that $\{\mathbf{v}_1\}$ is linearly independent, and we deduce that $\ell \geq 2$. Now, fix scalars $\alpha_1, \dots, \alpha_\ell \in \mathbb{F}$, not all zero, such that

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_\ell \mathbf{v}_\ell \quad = \quad \mathbf{0}.$$

By the minimality of $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$, we in fact have that $\alpha_1, \dots, \alpha_\ell$ are all non-zero.[10] We now multiply both sides of the equation above by the matrix A on the left, and we obtain

$$\alpha_1 A\mathbf{v}_1 + \dots + \alpha_\ell A\mathbf{v}_\ell \quad = \quad \mathbf{0}.$$

But note that for all $i \in \{1, \dots, \ell\}$, we have that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ (because $\mathbf{v}_i \in E_{\lambda_i}$). So,

$$\alpha_1 \lambda_1 \mathbf{v}_1 + \dots + \alpha_\ell \lambda_\ell \mathbf{v}_\ell \quad = \quad \mathbf{0}.$$

On the other hand, if we multiply our equation $\alpha_1 \mathbf{v}_1 + \dots + \alpha_\ell \mathbf{v}_\ell = \mathbf{0}$ by λ_ℓ instead, we obtain

$$\alpha_1 \lambda_\ell \mathbf{v}_1 + \dots + \alpha_\ell \lambda_\ell \mathbf{v}_\ell \quad = \quad \mathbf{0}.$$

So,

$$\alpha_1 \lambda_1 \mathbf{v}_1 + \dots + \alpha_\ell \lambda_\ell \mathbf{v}_\ell \quad = \quad \alpha_1 \lambda_\ell \mathbf{v}_1 + \dots + \alpha_\ell \lambda_\ell \mathbf{v}_\ell,$$

and consequently,

$$\alpha_1(\lambda_1 - \lambda_\ell)\mathbf{v}_1 + \dots + \alpha_{\ell-1}(\lambda_{\ell-1} - \lambda_\ell)\mathbf{v}_{\ell-1} \quad = \quad \mathbf{0}.$$

Since $\alpha_1, \dots, \alpha_{\ell-1}$ are all non-zero, and since $\lambda_1, \dots, \lambda_\ell$ are pairwise distinct, we see that the scalars $\alpha_1(\lambda_1 - \lambda_\ell), \dots, \alpha_{\ell-1}(\lambda_{\ell-1} - \lambda_\ell)$ are all non-zero. So, $\{\mathbf{v}_1, \dots, \mathbf{v}_{\ell-1}\}$ is linearly dependent, contrary to the minimality of $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$. This proves the Claim. ♦

Now, suppose that the set

$$\{\mathbf{v}_1^{\lambda_1}, \dots, \mathbf{v}_{t_1}^{\lambda_1}, \mathbf{v}_1^{\lambda_2}, \dots, \mathbf{v}_{t_2}^{\lambda_2}, \dots, \mathbf{v}_1^{\lambda_k}, \dots, \mathbf{v}_{t_k}^{\lambda_k}\}$$

[10] Indeed, if for some $i \in \{1, \dots, \ell\}$, we had that $\alpha_i = 0$, then $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\} \setminus \{\mathbf{v}_i\}$ would be linearly dependent, contrary to the minimality of $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$.

is linearly dependent. Then there exist scalars

$$\alpha_1^{\lambda_1}, \dots, \alpha_{t_1}^{\lambda_1}, \alpha_1^{\lambda_2}, \dots, \alpha_{t_2}^{\lambda_2}, \dots, \alpha_1^{\lambda_k}, \dots, \alpha_{t_k}^{\lambda_k} \in \mathbb{F},$$

not all zero, such that

$$\sum_{i=1}^{k} \sum_{j=1}^{t_i} \alpha_j^{\lambda_i} \mathbf{v}_j^{\lambda_i} \quad = \quad \mathbf{0}.$$

For each $i \in \{1, \dots, k\}$, set $\mathbf{v}_i := \sum_{j=1}^{t_i} \alpha_j^{\lambda_i} \mathbf{v}_j^{\lambda_i}$; then $\mathbf{v}_i \in E_{\lambda_i}$,[11] and moreover, since $\{\mathbf{v}_1^{\lambda_i}, \dots, \mathbf{v}_{t_i}^{\lambda_i}\}$ is linearly independent, we see that $\mathbf{v}_i = \mathbf{0}$ if and only if $\alpha_1^{\lambda_i} = \dots = \alpha_{t_i}^{\lambda_i} = 0$. Since not all $\alpha_j^{\lambda_i}$'s are zero, we see that at least one of $\mathbf{v}_1, \dots, \mathbf{v}_k$ is non-zero. After possibly permuting the order of our eigenvalues, we may assume that there exists some index $\ell \in \{1, \dots, k\}$ such that $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ are all non-zero, whereas $\mathbf{v}_{\ell+1} = \dots = \mathbf{v}_k = \mathbf{0}$. Then

$$\mathbf{v}_1 + \dots + \mathbf{v}_\ell \quad = \quad \mathbf{v}_1 + \dots + \mathbf{v}_k \quad = \quad \sum_{i=1}^{k} \sum_{j=1}^{t_i} \alpha_j^{\lambda_i} \mathbf{v}_j^{\lambda_i} \quad = \quad \mathbf{0}.$$

It follows that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is a linearly dependent set of vectors. But since $\mathbf{v}_1 \in E_{\lambda_1} \setminus \{\mathbf{0}\}, \dots, \mathbf{v}_\ell \in E_{\lambda_\ell} \setminus \{\mathbf{0}\}$, this contradicts the Claim.[12] □

Corollary 2.13. *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the following are equivalent:*

(i) $\mathbb{F}^n$ has a basis formed by eigenvectors of A;

(ii) the sum of algebraic multiplicities of all distinct eigenvalues A is equal to n, and the geometric multiplicity of each eigenvalue of A is equal to the algebraic multiplicity of that eigenvalue;

(iii) the sum of geometric multiplicities of all distinct eigenvalues A is equal to n.

Proof. Obviously, (ii) implies (iii). The fact that (iii) implies (ii) follows from the fact that the sum of algebraic multiplicities of the eigenvalues of A is at most n, and the fact that (by Theorem 2.5) the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

[11] Indeed, $\mathbf{v}_i$ is a linear combination of the basis vectors of E_{λ_i}, and so $\mathbf{v}_i \in E_{\lambda_i}$.

[12] One might object that, in order to apply the Claim, we need to have one non-zero vector out of each of $E_1, \dots, E_k$, not just out of $E_1, \dots, E_\ell$. However, this is easy to address: for each $i \in \{\ell + 1, \dots, k\}$, fix any eigenvector $\mathbf{u}_i$ of A associated with the eigenvalue λ_i; then $\mathbf{u}_i \in E_{\lambda_i} \setminus \{\mathbf{0}\}$. The Claim now guarantees that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell, \mathbf{u}_{\ell+1}, \dots, \mathbf{u}_k\}$ is linearly independent, and consequently, that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is linearly independent, too. But we showed that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is in fact linearly dependent! So, we do indeed have a contradiction.

It now suffices to show that (i) and (iii) are equivalent.

If (i) holds, then we see that the sum of geometric multiplicities of all distinct eigenvalues of A must be at least n;[13] since this sum cannot be greater than n,[14] we see that it is in fact equal to n, i.e. (iii) holds.

Suppose now that (iii) holds. Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of A. For each $i \in \{1, \dots, k\}$, let $\mathcal{B}_i$ be a basis for E_{λ_i}; by (iii), we have that $|\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k| = n$. On the other hand, by Theorem 2.12, $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ is a linearly independent set of vectors in $\mathbb{F}^n$. So, by Proposition 1.11(a) of Lecture Notes 7, $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ is in fact a basis of $\mathbb{F}^n$, and so (i) holds. $\square$

[13] Let us explain this in detail. Suppose that $\lambda_1, \dots, \lambda_k$ are the distinct eigenvalues of $\mathbb{F}^n$, and that $\mathcal{B}$ is a basis of $\mathbb{F}^n$ formed by eigenvectors of A. Then $\mathcal{B} \subseteq E_{\lambda_1} \cup \dots \cup E_{\lambda_k}$. Since $\mathcal{B}$ is linearly independent, it cannot contain more than $\dim(E_{\lambda_i})$ many vectors of E_{λ_i} for any $i \in \{1, \dots, k\}$; so, $|\mathcal{B}| \leq \dim(E_{\lambda_1}) + \dots \dim(E_{\lambda_k})$. On the other hand, we have that $|\mathcal{B}| = \dim(\mathbb{F}^n) = n$. So, $\dim(E_{\lambda_1}) + \dots \dim(E_{\lambda_k}) \geq |\mathcal{B}| = n$.

[14] This follows from Theorem 2.5, together with the fact that the sum of algebraic multiplicities of all distinct eigenvalues is at most n.